DOCUMENT RESUME

ED 172 736 IR 007 336

AUTHOR Willis, Barry
TITLE A Design Taxonomy Utilizing Ten Major Evaluation Strategies.
PUB DATE 27 Dec 79
NOTE 26p.; Not available in paper copy due to marginal legibility
AVAILABLE FROM Dept. of Instructional Media, Utah State University, UMC 30, Logan, Utah 84322

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Educational Methods; *Evaluation Criteria; *Evaluation Methods; Evaluation Needs; Formative Evaluation; Summative Evaluation; *Taxonomy

ABSTRACT

This paper discusses ten evaluation strategies selected on the basis of their general acceptance and their relatively unique approach to the field: (1) Stake, "Countenance of Evaluation"; (2) Stufflebeam, "Decision Centered Evaluation (CIPP)"; (3) Provus, "Discrepancy Evaluation"; (4) Scriven, "Goal Free Evaluation"; (5) Scriven, "Formative and Summative Evaluation"; (6) Blaha, "Trade-Off and Comparative Cost Approach"; (7) Hunter and Schooley, "The Synergistic Evaluation Strategy"; (8) Foster, "The Satisfactory Evaluation Strategy"; (9) Crowl, "The Crowl Evaluation Strategy"; and (10) Stufflebeam/Scriven/Guba, "Meta-Evaluation, or the Evaluation of Evaluations." Educational evaluators are advised to systematically pick and choose among the elements composing the various evaluation strategies in order to find the combination of elements appropriate for a given evaluative need. A taxonomy of the strategies discussed is included, as well as a list of references. (JTG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED172736

IR

(FINAL DRAFT)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A DESIGN TAXONOMY UTILIZING TEN MAJOR EVALUATION STRATEGIES

BEST COPY AVAILABLE

December 27, 1978

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Barry Willis

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Prepared by:

Barry Willis, Director of
Instructional Product Development
Center for Instructional Product
Development
Utah State University
Logan, Utah

Send reprint requests to:
Barry Willis
Dept. of Instructional Media
Utah State University UMC 30
Logan, Utah 84322

IR 007336

ABSTRACT

When analyzing evaluation strategies, one must systematically pick and choose among the elements composing the various approaches. While completing this task, it is important to keep in mind that the purpose of an evaluation strategy is to aid in evaluative decision making. By utilizing a taxonomy approach, it is possible to analyze and compare various strategies in efforts to pick the elements needed to satisfy a given evaluative need.

INTRODUCTION

We all have our problems. So it is with educational evaluators. Wandering the streets of academia in search of an identity, educational evaluators are caught in a conflict arising from a sense of not belonging.

Whether due to fear, frustration, or just a general lack of understanding, evaluation has never, until just recently, been given more than cursory acceptance.

Stufflebeam (1971) reported, "Avoidance of evaluation is evidenced even within the United States Office of Education (USOE) which, historically, has provided neither budget or staff sufficient to evaluate its own programs, despite a great deal of talk about the desirability of evaluation for the school." [1] As in any field where strict definitions are neither appropriate nor possible, evaluation begs the participant to constantly reassess his options in search of a better way.

The problem arises when one tries to differentiate among these better ways. In the field of evaluating programs for disadvantaged adults alone, the ERIC Clearinghouse on Adult Education lists 58 evaluation models currently in use.

By closely scrutinizing these as well as a plethora of other strategies, one finds a pattern of similarities if not total repetition. In viewing these various evaluative approaches, it becomes apparent that evaluators generally follow one of several basic patterns and then embellish the similarities with a personal touch in efforts to legitimize their "new creation." This approach typifies, in many respects, the expanding field of educational technology in general and instructional development in particular. As in evaluation, instructional developers have settled on a

basic approach, focusing on educational design, development, evaluation, and revisions, prior to massaging additional steps geared to reflect their unique circumstances.

The same generalization holds for evaluation--there is a basic structure that is embellished to meet one's unique needs. The question arises: Is it possible to isolate particular evaluative characteristics and then build, in cookbook fashion, a super-strategy based on the various major evaluative approaches? In essence, that is the purpose of this paper--to analyze the major program evaluation strategies utilized by the current leaders in the field by developing a taxonomy of evaluative strategies.

While delving into this analysis of evaluation strategies, heed a word of caution voiced by W. James Popham, a seasoned evaluation specialist from the University of California at Los Angeles (UCLA): "Although it is sensible for educational evaluators to inform themselves of the nature of educational evaluation models, they should not get too caught up in that enticing but enervating game known as comparative model meshing. Some people take great delight in seeing how Model X differs from Model Q and is ever so slightly like the seventy-ninth stage of Model Z. Instead of engaging in a game of "sames and differents," the educational evaluator should become sufficiently conversant with the available models of evaluation to decide which, if any, to employ. Often a more eclectic approach will be adopted whereby one selectively draws from several available models those procedures or constructs that appear most helpful." [2]

WHY A TAXONOMY?

For the purposes of this paper, "taxonomy" is defined as a systematic attempt at classifying phenomena into categories, either corresponding to real ordering or to arbitrary ordering among the phenomena.

Concerning the development of an evaluation design taxonomy, Worthen (1968) states, "While manageable in theory, such an approach presents the evaluator with a task which is extremely difficult in practice. It requires that the selection of alternatives affects a balance between selection of the most appropriate alternative for each decision situation and the selection of a set of alternatives which can be integrated into the best over-all design." [3]

In analyzing various evaluation designs, Stufflebeam (1978) developed a logical evaluation design structure that transcends all types of evaluation, whether content, input, process, or product--the components of his own model. [4] The necessary elements are as follows:

A. Focusing the Evaluation
   1. Identify the major level(s) of decision-making to be served e.g., local, state, or national.
   2. For each level of decision-making, project the decision situations to be served and describe each one in terms of its locus, focus, timing, and composition of alternatives.
   3. Define criteria for each decision situation by specifying variables for measurement and standards for use in the judgment of alternatives.
   4. Define policies within which the evaluation must operate.

B. Collection of Information
   1. Specify the source of the information to be collected.
   2. Specify the instruments and methods for collecting the needed information.
   3. Specify the sampling procedure to be employed.
   4. Specify the conditions and schedule for information collection.

C. Organization of Information
   1. Specify a format for the information which is to be collected.
   2. Specify a means for coding, organizing, storing, and retrieving information.

D. Analysis of Information
   1. Specify the analytical procedures to be employed.
   2. Specify a means for coding, organizing, storing, and retrieving information.

E. Reporting of Information
   1. Define the audiences for the evaluation reports.
   2. Specify means for providing information to the audiences.
   3. Specify the format for evaluation reports and/or reporting sessions.
   4. Schedule the reporting of information.

F. Administration of the Evaluation
   1. Summarize the evaluation schedule
   2. Define staff and resource requirements and plans for meeting these requirements.
   3. Specify means for meeting policy requirements for conduct of the evaluation.
   4. Evaluate the potential of the evaluation design for providing information which is valid, reliable, credible, timely, and pervasive.
   5. Specify and schedule means for periodic updating of the evaluation design.
   6. Provide a budget for the total evaluation program.

These criteria along with others developed by the author will be utilized in development of this matrix-based evaluation taxonomy.

TEN EVALUATION STRATEGIES FOR COMPARISON

For the purposes of this taxonomy, the following evaluation strategies were selected on the basis of their major acceptance and their relatively unique approach to this complete field. The individual strategies along with a brief annotation of each are provided below:

1. STATE--"Countenance of Evaluation"

Evaluation data are either descriptive--intents and observations--or judgmental. In order to evaluate, an educator will gather together certain data. The data are likely to be from several quite different sources, gathered in several quite different ways. Whether the immediate purpose is description or judgment, three bodies of information

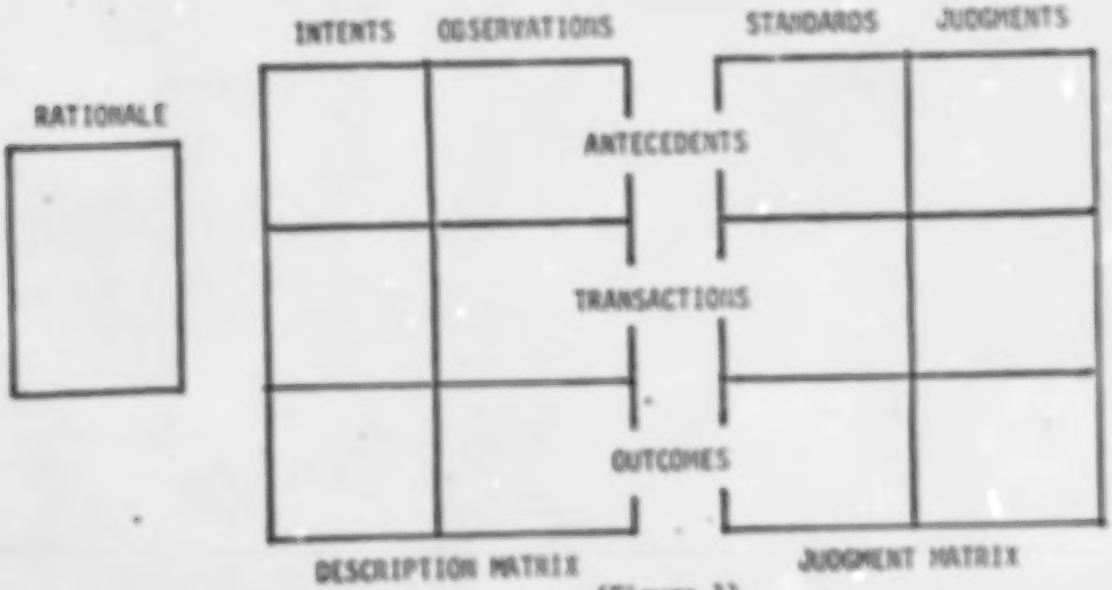
should be tapped. In the evaluation report it can be helpful to distinguish among antecedent, transaction, and outcome data.

An antecedent is any condition existing prior to teaching and learning which may relate to outcomes. The status of a student prior to his lesson, e.g., his aptitude, previous experience, interest, and willingness, may comprise a complex antecedent.

Transactions are the countless encounters of students with teachers, students with student, author with reader, etc. Examples include the presentation of a film, a class discussion, etc. Transactions are dynamic whereas antecedents and outcomes are relatively static.

Outcomes are the consequences of educating--immediate and long-range, cognitive and conative, personal and community-wide.

Judgmental statements are classified either as general standards of quality or as judgments specific to a given program. Descriptive data are classified as intents and observations. The evaluator can organize his data gathering to conform to the format shown in Figure 1. [5]



(Figure 1)

The Stake model is considered to be the most humanistic of any of the major models in current use, placing heavy weight on judgmental data.

2. STUFFLEBEAM; PHI DELTA KAPPA--"Decision-Centered Evaluation (CIPP)"

Evaluation provides information for judging decision alternatives. It can be useful at all stages of decision making--awareness, design, choice, and action. Four levels of evaluation--context, input, process, and product--correspond to four major steps of programming decisions--planning, structuring, implementing, and recycling.

3. PROVUS--"Discrepancy Evaluation"

Evaluation identifies discrepancies between actual programs and standards for programs so that programs can be improved. Program personnel set standards for activities and results at each programming stage--design, installation, process, and product. Actual performance is compared with the standard and discrepancies or areas for improvement identified. Discrepancy information is used either to change the performance or to change the standard.

4. SCRIVEN--"Goal Free Evaluation"

Results of programs are judged against the originating need rather than the stated objectives. Cost of producing those results is considered in relation to costs of alternatives. The work of evaluation is viewed as primarily an act of condensation that includes two major stages--compression and credentialling. A mass of data and observations are compressed until judgments as to the value or worth of the program can be made--credentials assigned. [6]

5. SCRIVEN--"Formative and Summative Evaluation"

There are two evaluative stages: formative--the purpose of which is to assist in developing curricula--and summative--the purpose of which is to assess the merit of curricula once they have been developed and are on the market. Scriven states that formative evaluation is an ongoing process providing constant feedback to the producer with summative evaluation being used to assess the merit of the finished product by the consumer.

6. GLASS--"Trade-off and Comparative Cost Approach"

A format for evaluation of materials or educational activities, including: describing the product to be evaluated; evaluating the goals of the product; clarifying the point of entry of the evaluator; determining the kinds of trade-offs involved; comparing costs with costs of alternatives; making an intrinsic (secondary) evaluation; making an outcome (primary) evaluation; forming judgments and recommendations; stipulating circumstances that would modify the conclusions; and evaluating the evaluator.

7. HUNTER AND SCHOOLEY--"The Synergistic Evaluation Strategy"

An educational system can be broken down into four domains: (1) policy, (2) program development, (3) instruction, and (4) feedback. Each domain contains activities which are of a technical and non-technical nature which comprise a complete evaluation strategy. The approach is synergistic in three ways: (1) its activities require the involvement and complete interaction among students, educators, and parents; (2) it requires the interaction between technical and non-technical aspects of evaluation;

and (3) its activities include both goal based evaluation (GBE) and goal free evaluation (GFE). [7]

8. FOSTER--"The Karlsruhe Evaluation Strategy"

As an offshoot of the Scriven Formative/Summative evaluation system, the basic steps involved are: (1) preplanning, (2) planning and model development, (3) implementation, (4) operation, and (5) dissemination. In contrast to most developmental strategies, evaluation is not listed as the final phase because evaluation is viewed as an ongoing activity at all stages of program development.

9. CRIDEL--"The Cremel Evaluation Strategy"

The Cremel strategy is developed utilizing an X/Y matrix. The X axis is concerned with the following categories: (1) issues, (2) initiation, (3) pilot, (4) field test, and (5) public diffusion. The Y axis is concerned with: (1) desirability/feasibility, (2) management procedural cost, (3) product worth, (4) usability, and (5) generalizability. The evaluator responds to the matrix in terms of: (1) criteria, and (2) audience. This enables those involved in the study to analyze the individual segments of the evaluation as to their instructional worth.

In using the Cremel strategy, the evaluator helps the client to identify the issues to be considered, i.e., "Desirability/Feasibility." The evaluator/client team then identifies the various audiences for each issue (varies from issue to issue) and the criteria needed to satisfy a particular audience in regards to a specific issue. Similar criteria/audience analysis is conducted for the various stages of program evaluation, i.e., initiation, pilot, field testing, public diffusion. (Figure 2)

(Fig. 7)

a. What are the worst needs of the university?										
b. Is this problem a high priority issue?										
c. Are there products available or in development which might be adapted to this purpose?										
d. What type of product is necessary to fill the need?										
e. Identification of target population.										

10. STUFFLEBEAM/SCRIVEN/GLASS--"Meta-Evaluation, or, the Evaluation of Evaluations"

Because of its varied interpretations, a little background information on Meta-Evaluation is in order. The term "Meta-Evaluation" means "A procedure for describing an evaluation activity and comparing it against a set of ideas concerning what constitutes good evaluation." According to Scriven, this means that meta-evaluation is higher order than standard evaluation and includes evaluations that are secondary, tertiary, etc. [8]

Criteria for judging meta-evaluation have been discussed in the works of Campbell and Stanley [9]; Gephart, Ingle and Reinstad [10]; and Bracht and Glass. [11]

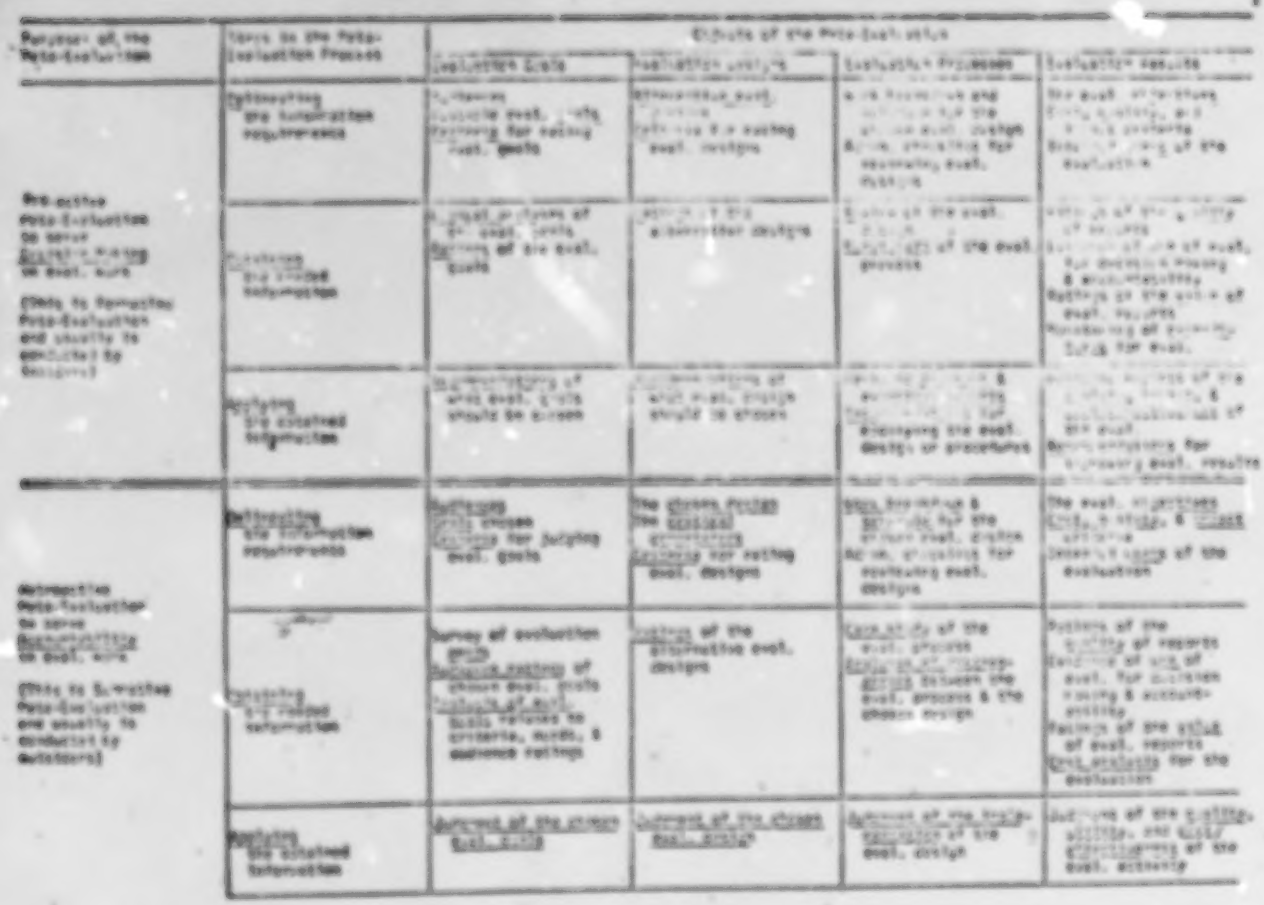


Figure 3
A Meta-Evaluation Framework

The structuring of meta-evaluation appears in Figure 3. This structure portrays meta-evaluation as a methodology for assessing the merit of proposed and completed evaluation efforts as discussed in the first premise. The framework relates to three dimensions, namely; PURPOSES, OBJECTS, AND STEPS. The contents of the cells of the structure reflect the three criteria of technical adequacy, utility and cost/effectiveness. Meta-evaluation follows the premise that insiders should conduct formative or pro-active meta-evaluation while external agents should conduct summative or retroactive meta-evaluation.

Stufflebeam states that there are two specific purposes of meta-evaluation: [12]

1. Meta-evaluation should serve decision making and accountability. In order to support decision making in evaluation, it requires and emphasizes the need for formative work in efforts to provide timely recommendations concerning how evaluation studies should be designed and carried out. Formative meta-evaluation is a direct way of insuring that evaluations will produce results that are technically adequate, useful, and cost/effective.

2. The second, and possibly outwardly more important, purpose of meta-evaluation is to provide a system of accountability for the evaluator. To facilitate this accountability, meta-evaluation of a summative type is conducted by outside agent to produce an un-biased outside opinion and foster public judgments of the merits of the completed evaluation work. Much of the information required in summative meta-evaluation is potentially available in the formative meta-evaluation.

MATRIX FORMAT

In assessing the various evaluation strategies, several characteristics were deemed necessary to explore. They are:

A. Type of Evaluation--The context in which the evaluation is conducted.

B. Definition--The primary aim of the evaluation.

C. Purpose--What the results of the evaluation will be used for.

D. Criteria for Judging Evaluation--The criteria by which the evaluative design and implementation will be judged.

E. Personnel Required--Staffing implications in relation to a given evaluative strategy, i.e., to successfully meet evaluative criteria certain strategies need more manpower than others. (This is an important budgetary consideration.)

F. Implications for Design--Specific design constraints encountered when following a given strategy.

G. Key Emphasis--How the information gleaned from the evaluation is used.

H. Role of Evaluator--The evaluator's role in regards to a specific strategy.

I. Relationship to Objectives--Means by which the objectives are met.

J. Relationship to Decision Making--How the evaluation strategy serves decision making.

K. Time to Complete--Approximate time needed to complete a given evaluation. Time varies greatly concerning the size of the program and depth of the evaluation.

L. Cost--Budgetary requirements in very general terms dependent, once again, on the size of the program and depth of the evaluation.

It is quite obvious that various other characteristics could be plugged into this framework. It is felt, however, that the characteristics listed here will provide the reader with a firm base from which various evaluation strategies can be analyzed.

LIMITING FACTORS

Every taxonomy has inherent drawbacks. While some are quite obvious, others remain hidden for the most part.

In looking over the characteristics considered in the matrix, it is apparent that other aspects of the models could have easily been added to this framework. While additional characteristics would make the taxonomy more comprehensive, they would also tend to make the taxonomy more difficult to digest. Another problem arises from the difficulty of classifying various strategies using somewhat vague terminology such as "cost,"

"purpose," etc. Terminology that is too rigid will tend to misrepresent certain strategies while non-specific terms will fail to give a firm base by which different strategies can be analyzed.

A final factor of concern to the person actually conducting an evaluation is the size of the program being analyzed. One should keep in mind that the size of an evaluative effort, both in terms of budget, personnel requirements, deadlines, etc., will have a great affect on the evaluative strategy or strategies to be utilized.

[illegible] Evaluation Theorist	STAKE Countenance Strategy	STUFFLEBEAM CIPP	PROVUS Discrepancy Evaluation	SCRIVEN Goal Free	SCRIVEN Formative/Summative	HAMMOND Trade-Off and Comparative Cost Approach	HENTER and SCRIBNER Synergistic Evaluation Strategy	POTTER Bartender Strategy	CRONBACH Client Strategy	[illegible] Non-Evaluation
TYPE OF EVALUATION	(1) Formal vs. informal utilizing a heuristic approach	(1) Context (2) Input (3) Process (4) Product	(1) Design (2) Installation (3) Process (4) Product (5) Cost	formative/Summative in nature lending more weight to program "side effects" and "main terminal results" than goal based evaluation that focuses only on a project's ability to meet established goals	(1) formative/summative (2) requirements/consequitive (3) intrinsic/payoff (4) Mediated	A form of Goal based evaluation that is very decision oriented and extremely well documented throughout the entire process	Based on interaction between technical and non technical aspects of evaluation using both goal based evaluation (GBE) and goal free evaluation (GFE) [14] Includes formative and summative elements	A program-wide evaluation system formative and summative in nature. Evaluation instruments are developed for each phase of the project focusing on the following major program phases: (1) preplanning (2) planning and model development (3) implementation (4) operation (5) dissemination	Formative/Summative design utilizing an "issue" based matrix to determine alternatives to instructional situations	Formative/Summative [illegible] to serve [illegible] making [illegible] judgments in [illegible] and accountability [illegible]
DESCRIPTION	Describing and judging an educational program utilizing data matrices concerning antecedents, transactions and outcome data. (See fig. 1)	Defining, obtaining and using information for decision making	Comparing performance against standards with the difference being classified as the discrepancy	Goal-free evaluation is concerned more with actual outcomes vs. intended outcomes. GFE is unstructured in that it gives no real guidance in gathering data. It is left up to the goal-free evaluator to choose among the thousands of potentially relevant attainment variables and associated measuring devices. GFE does not provide specific guidance in choosing among them	Gathering and combining performance data with weighted set of goal scales	The goal of an evaluation is to describe the program outcomes and to discuss evaluator recommendations. This can only be done with the preliminary establishment of program goals. The result of such an evaluation will be an analysis of trade-offs in the form of recommendations which might better help a project attain its stated goals	An educational evaluation strategy designed to create closer interaction between evaluation researchers and curriculum specialists. The model separates an educational system into four domains: (1) policy (2) program development (3) instruction (4) feedback	This strategy formalizes the developer's intuitive process for establishing program goals by formalizing that intuition and combining the relevance of the developer's approach with the explicitness and rigor of the scientific approach. It focuses on the 5 steps described above [15]	Any evaluation is dependent on certain identified or identifiable situation. By analyzing a situation in terms of (1) resources (2) contexts and audience, a given program can be analyzed in terms of initiation (development), pilot testing and field testing	"A procedure for describing an evaluation activity and comparing it against a set of ideas concerning what constitutes good evaluation." [6] In other words, the evaluation of evaluation
PURPOSE	To describe and judge educational programs based on a formal inquiring process	To provide relevant information to decision makers	To determine whether to improve, maintain, or terminate a program	The purpose of GFE is to analyze the actual effects a product or program has, whether or not they were intended [16]	To establish and justify merit or worth of a given project by using evaluation in all phases of program development	To analyze a program in relation to its goals and to explore trade-offs and budgetary options which will enable a project to meet its primary goals within its budget requirements	The purpose of this strategy is to analyze a school system by breaking it down into the six groups comprising the educational system: (1) students (2) instructional staff (3) administration staff (4) educational specialists (5) family (6) community	To assess the legitimacy of an organization's developmental process through individualized evaluative instrumentation	To analyze pertinent program issues at the start of a project. The efforts to evaluate possible problems and situations before they occur	(1) To serve accountability and decision making (2) To form a system of accountability for the evaluator

	(1) Should be pervasive, not micro-scopic (2) Should include descriptive and judgmental data (3) Should provide to evaluate relative answers for decision making (4) Should be formal (e.g., objective, scientific, reliable)	(1) Internal validity (2) External validity (3) Reliability (4) Objectivity (5) Relevance (6) Importance (7) Scope (8) Credibility (9) Timeliness (10) Pervasiveness (11) Efficiency	(1) Team involvement (2) Assures one-to-one correspondence between design and solution (3) Compare performance against standards as a tool for improvement and assessment (4) Periodic feedback	There are no criteria for judging the evaluation other than actual results, where an ideal evaluation predicts test actual results vs. intended results	(1) Should be predicated on goals (2) Must indicate worth (3) Should have construct validity (4) Should be a holistic program evaluation	The conclusion must be well documented with various trade-offs expanded to the point where options and their consequences can be analyzed by decision makers	Goals and objectives are developed for the four domains. It is synergistic in that involvement and input are gathered from students, professional educators, and parents	Through the use of specialized evaluation instruments focusing on specified program phases, the evaluator can assess the program's progress in terms of outcome as well as systematically developed inputs such as pre-planning, goal analysis, and selection of a suitable development strategy	The is the audience for the various aspects of a given program? What criteria must be met to fill these needs?	(1) Who is to be served? (2) What are their needs? (3) How does the audience want to learn? (4) What action will be taken upon completion?
	The fewer, the better, with one-on-one interaction the ideal. Although formal, it is very responsive to individual need and purpose	1 grad team evaluation with one head evaluator and 4 other team members covering each of Stufflebeam's components: (1) context (2) input (3) process (4) product	A team of 3 to 5 members is typical with numbers of evaluators working on a given project ranging up to 10, depending on time constraints	In a given project, one goal from evaluators should work independently, beginning midway in the point after completion. While this will not guarantee good quality and fair evaluation, it will provide an opportunity to eliminate the "error term" contend to [illegible]	Works well with two teams--an internal team concerned with formative evaluation and an external team concerned with summative evaluation	Best handled as a team effort with members critiquing various trade-offs	Utilizes a team of 4 to 7 people	Best handled as a team composed of 4 to 7 evaluators. Each evaluator is assigned a project phase and has the responsibility to develop the instrumentation for assessing the program's ability to meet the requirement of that phase	After project staff input is utilized and issues agreed upon, a single evaluator could guide the evaluation process	Formative evaluation can be conducted internally and monitored by the external evaluation team leader who oversees the completion of the summative evaluations
	Four general structure utilizing two design matrices--one for "description" and the other for "judgment." Each matrix is concerned with antecedents, transactions, outcomes in point, and intents, observations (descriptive matrix), and standards, judgments (judgment matrix)	(1) Experimental design not applicable (2) Use of systems approach for evaluation studies (3) Directed by administrator	(1) Provide continuous evaluation through feedback loops (2) Provide relevant & timely information for decision making (3) Provide cost benefit analysis (4) Involvement of evaluator in program development	Unstructured to the degree that the individual evaluator selects the variables and relevant measuring devices based on intuition	(1) Look at many factors (2) Be involved in value judgments (3) Require use of scientific investigation (4) Evaluate from within (formative) and from without (summative)	Goal based focusing on: (1) product description (2) evaluation of goals (3) clarification of evaluator's point of entry (4) determination of trade-offs (5) examining cost alternatives (6) internal/evaluation/outcome (7) judgments, recommendations (8) circumstances that could modify conclusions (9) meta-evaluation	Representatives of the six groups should be identified early as the project interaction between the researchers and curriculum developers must be maintained throughout processing of all four domains	Evaluation instrumentation is project specific in that the evaluation team must develop specific instruments to assess the quality and worth of an organization's systematic approach. Although instrumentation varies, the same basic model outline is still one to follow	Fairly simple structure. Major areas headings such as: (1) usability, (2) desirability of feasibility, and (3) management/ procedural/cost are provided with specific outlined things dealing with the particular project. Issues are dealt with in terms of (1) science, and (2) criteria	Focus on the four aspects of metaevaluation in a formative sense for the active (1) evaluation goals, (2) evaluation designs, (3) evaluation procedures, (4) evaluation results, then arrive at a retrospective assess the overall merit and of the evaluation
	Collection of descriptive and judgmental data from various audiences to aid in decision making	Evaluation reports used for decision making	Identifying discrepancies between standards and performance using a team approach	Evaluators should be concerned with actual program outcomes rather than predetermined goal	Identification of data gathering, instrumentation, weightings, and selection of goals. Evaluation strategy combines data on different performance scales into a single rating	Identifying trade-offs and cost alternatives for decision making purposes	Gathering interactive data on group selected goals and objectives in an effort to evaluate a program utilizing all parties involved in the development and implementation of an educational system	Assessing the organization's development model in terms of its ability to systematically attain identified goals	Once specific issues are agreed upon and established, the evaluator can analyze a program in terms of the ability to meet the program needs as stated in the issues	Evaluating the results of evaluations using formative and summative data

VII Evaluation Strategy	STAKE Countenance Strategy	STUFFLEBEAM CIPP	PROVUS Discrepancy Evaluation	SCRIVEN Goal Free	SCRIVEN Formative/Summative	BLAKE Trade-Off and Comparative Cost Approach	CARTER and ROSSMILLET Synergistic Evaluation Strategy	HOUSE Nationale Model	CRONBACH Caronet Strategy	GUBA Meta Exploration
ROLE OF EVALUATOR	Specialist concerned with collecting, processing, and interpreting descriptive & judgmental data	Specialist who provides evaluation information to decision making	A team member who aids program improvement and consults administration. The team member should be independent of the program unit	Goes to actual program outcomes, unencumbered and unbiased by preconceptions in regards to program goals	Responsible for judging the merit of an educational practice for producers (formative) and consumers (summative)	To provide information on project options to decision makers as well as provide summative information	To spark intergroup communication, to set up & monitor the 4 domains, to make sure all relevant groups are represented and to conduct summative evaluation in the feedback stage	The evaluator role is value-free in that the organization's development process is analyzed and described by the evaluator but not judged	(1) To speed up the development of the individual in-uses of a given program (2) to mobilize behalf the issues are eloquently addressed to project initiation (brainstorming), pilot testing, field testing, and public deliberation	To evaluation must [illegible] a true part of all stages of [illegible] evaluation i.e., formative evaluation [illegible] design are needed [illegible] The evaluation as the central [illegible] part of the meta-evaluation study
RELATIONSHIP TO OBJECTIVES	Examination of goal specifications and prioritizes. Identification of areas of failures and successes. It is up to the evaluator to assist in writing behavioral objectives	The terminal stage in context evaluation is setting objectives; input evaluation produces ways to reach objectives; product evaluation determines whether objectives are reached	Agreement of evaluation team & program staff on evaluation criteria. Comparison of performance against standards to see whether a discrepancy exists [13]	No relationship whatsoever	Look at goals and judge their worth. Determine whether they are being met	A direct relationship to goals and objectives in that the entire evaluation is based on the assessed viability of program goals	Goals & objectives are developed interactively by the six groups composing the educational system. In the final task domain (#6), attainment of the goals and objectives are measured	Objectives for each phase are specified by the organization with the evaluator's role one of developing instruments to assess the organization's progress and progress in meeting those specified objectives	Objectives are developed in terms of program "issues" & addressed in terms of criteria, in each of the 5 program stages: (1) initiation (brainstorming) (2) pilot testing (3) field testing (4) public deliberation	Evaluation goals are developed throughout the process, results of which are discussed & placed in the evaluations report
RELATIONSHIP TO DECISION MAKING	Descriptive and judgmental data in reports (including recommendations to various audiences). Judgments may be based on either absolute or relative standards	Evaluation provides information for use in decision making	Evaluation staff collects information essential to program improvement & notes discrepancies between performance and standards. Every operation involves a criterion (C), new information (I), & a decision (D). Evaluation provides new information for the decision maker	Vague, in that actual results are documented, but not necessarily compared to intended results (goals)	Evaluation reports (with judgments explicitly for producers & consumers) used in decision making	Decision makers receive program trade-offs & alternative cost approaches which they use to modify the program to better meet its goals	In the feedback domain (#4) summative data is gathered based on accomplishments of goals and objectives. All summative material is documented in a final report	A straight forward report is developed by the evaluation team (value free) describing the data gathered using the various evaluation instrumentation	The data is gathered, documented, analyzed and interpreted into a report prior to being used by decision makers. Decisions based on decision making are: (1) [illegible] (2) [illegible] (3) product worth (4) [illegible] (5) generalizability	All material gathered in the various phases of the countenance and transaction evaluations are retained & documented in the meta evaluation report. The report is then used by decision makers
TIME TO COMPLETE	Can be varied depending on needs. Time is needed to develop trust level time varies depending on quantity of people to be queried & openness of those interviewed	Could be done relatively quickly--using the evaluation team concept	Several months of initial/about team effort then 2 to 3 week of analysis & development of new information (I)	Since it is both formative and summative in nature, it would run the length of the project	Formative is ongoing during program development and lasts the length of design & development--summative is handled in a field test	Takes quite a bit of time to analyze a given program and develop trade-off and cost alternatives	Depends largely on the team members representing the different factions. The shorter more intense effort, the better	Since this form of evaluation is ongoing-input gathered throughout the program, the length of time required to complete the project is dependent on program duration	Since the process is formative and summative in nature, the completion time is dependent on the duration of the project	[illegible] with project initiation, [illegible] the length of the program [illegible] and is not feasible until interactive instruction data is gathered, documented, and [illegible]
COST	Can be fairly expensive depending on depth and breadth of the evaluation	Relatively inexpensive if the team concept is utilized. The team leader could oversee the 4 less experienced evaluators and oversee development of the various evaluation reports	Expensive, ranging from 10 to 15 percent of the program cost	Varies depending on goals invented. A goal-free evaluation can be adequately accomplished by one evaluator. Some suggest that 2 independent evaluators should work on a project simultaneously although this would double the cost per man day it would likely still cost less to complete a GFE than many of the other forms of evaluations	Varies. Formative can be partially by program staff for a minimal cost with summative complimented by an external staff usually costing more	Fairly expensive. It takes a very experienced evaluator to be able to analyze options and come up with feasible trade-offs and alternative cost approaches	Comparatively inexpensive since all of those involved with the possible exception of the primary evaluator and possibly an assistant or two are volunteers representing the six primary groups	Fairly expensive since new instrumentation is needed for each phase of a given program	Fairly inexpensive. Roughly 3 to 5 man days to set up, several more for monitoring purposes and approximately to supervise final data and develop a concluding report	Meta-evaluation can and should be fairly complex and for this reason is relatively costly

SYNTHESIS

As stated earlier in this paper, the vast majority of evaluation strategies utilizes similar events in a rather predictable sequence. In analyzing the various strategies discussed in this paper, it is noteworthy that nearly all frameworks utilize elements of existing strategies to come up with "unique" evaluative approaches. For example, most strategies follow a sequence similar to Stufflebeam's Context, Input, Process, Product (CIPP) system for laying out the different evaluative stages. Almost all models include the formative/summative evaluative process credited to Scriven. Although Scriven popularized the term "formative/summative evaluation," the vast majority of present day evaluative strategies before that time utilized some form of program feedback for guidance from project initiation to completion.

Other approaches, such as Stake's "Countenance of Evaluation" system attempt to gather similar data utilizing different methods; such as Stake's humanistic approach to gathering information.

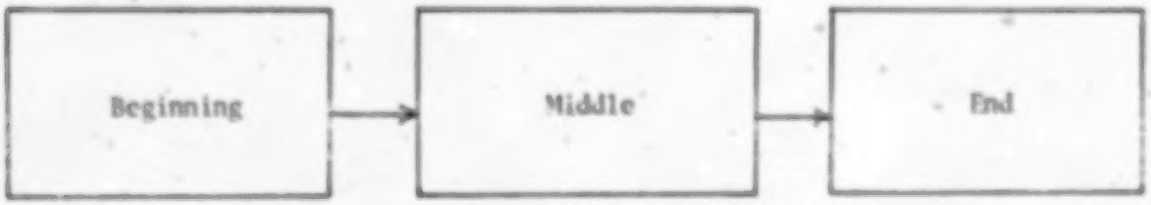
An example of evaluating a given product or program using a unique approach is Scriven's "Goal-Free" evaluation strategy in which goals, the guiding light of most evaluators, are intentionally by-passed in efforts to create an outcome-based, bias-free evaluation.

Early in the development of this taxonomy, the need was expressed to find a "super strategy" utilizing the best elements of each evaluative approach. In analyzing the approaches discussed in this paper, it becomes obvious that the development of a "super strategy" for educational evaluation is not a new concept . . .

. . . Each of the strategies discussed in this paper is an attempt at developing the ultimate evaluation tool designed to fill a specific need. The

success of various strategists in reaching this end can only be decided by the individual evaluator, who, after assessing his evaluation needs, must look at evaluation options currently available and decide whether to utilize an existing one or develop a "super strategy" of his own which transcends previous models.

One needs to remember, however, that evaluation strategies need not be complex to be effective. There is, for example, a particularly simple one developed by Ernest R. House, and reported in his article: "Confessions of a Responsive Goal-Free Evaluator." Cloaked in the jargon of educational evaluation, the tongue-in-cheek House states, "Below is an evaluation model developed by an Office of Education official and reported in Educational Technology (1963).



Or was it an evaluation model of the Office of Education? In any case, the data will be arrayed into the above scheme." [18]

House's point is well taken. The purpose of an educational evaluation strategy should be to facilitate evaluative decision making. This should be done as efficiently as possible without undue glorification and obfuscating frills. While House's simplistic model overstates the case, the measure of a "good" evaluation strategy is its ability to fill an expressed evaluative need and not to add more jargon to a field approaching the saturation point.

LITERATURE CITED

1. Stufflebeam, D.L. "Educational Evaluation and Decision Making," Itasca, Ill, F.E. Peacock Pub, 1971

2. Popham, W. James "Educational Evaluation," Prentice-Hall, Inc p. 21, 1975

3. Worthen, Blaine "Toward a Taxonomy of Evaluation Designs," Educational Technology, August 15, 1978

4. Stufflebeam, D.L. "Toward a Science of Educational Evaluation," Educational Technology, July 30, 1968

5. Stake, Robert E. "The Countenance of Educational Evaluation," Teachers College Record, Vol. 68, No. 7, April 1967

6. Steele, Sara "Contemporary Approaches to Program Evaluation: Implications for Evaluating Programs for Disadvantaged Adults," Capital Pub, 1976

7. Alkin, M. "Wider Context Goals and Goal-Based Evaluators," The Journal of Educational Evaluation: Evaluation Comment, Vol. 3, No. 4, 1972

8. Scriven, Michael S. "An Introduction to Meta-Evaluation," Educational Product Report, Vol. 2, No. 5, Feb. 1969

9. Campbell, Donald T. and Julian C. Stanley "Experimental and Quasi-Experimental Designs for Research on Teaching," Handbook of Research on Teaching, Chicago, Ill, Rand McNally, 1963

10. Gephart, W.J. and R.B. Ingle and R.C. Remstad "A Framework for Evaluating Comparative Studies. Conference on Research in Music Education," U.S. Office of Education Cooperative Research Report No. 6, May 1967

11. Bracht, Glenn H. and Gene V. Glass "The External Validity of Experiments," American Educational Research Journal, Nov 1968

12. Stufflebeam, Daniel L. "Meta-Evaluation," Evaluation Center, Occasional Paper, No. 3, 1977

13. Anonymous. Guide to Using the Discrepancy Evaluation Model for the Staff of Projects Supported by the Developmental Disabilities Council, Exceptional Child Center, 1976

14. Alkin, M. "Wider Context Goals and Goal Based Evaluators," The Journal of Educational Evaluation: Evaluation Comment, Vol. 3, No. 4, 1972

15. DeBlois, Michael, "An Evaluation Paradigm for Flexible Staffing Patterns and its Application to the Temple City, Mass. and Florida Network Projects," The Evaluation Training Center, Tallahassee, Florida, Vol. 3, 1971

16. Scriven, Michael, "Prose and Cons About Goal-Free Evaluation," The Journal of Educational Evaluation: Evaluation Comment, Vol. 3, No. 4, 1972

17. Stufflebeam, Daniel L. "Should or Can Evaluation Be Goal-Free?" The Journal of Educational Evaluation: Evaluation Comment, Vol. 3, No. 4, 1972

18. House, Ernest R. "Confessions of a Responsive Goal-Free Evaluator." Center for Instructional Research and Curriculum Evaluation Occasional Paper, University of Illinois at Urbana-Champaign, 1974